Project: UVMR13-7

DERIVING LAND-USE LIMITS AS A FUNCTION OF INFRASTRUCTURE CAPACITY

Final Report

by

Adel W. Sadek, Assistant Professor
Department of Civil and Environmental Engineering
University of Vermont
Burlington, VT 05405
E-mail: asadek@emba.uvm.edu

Wael M. ElDessouki, Assistant Research Professor
Department of Civil Engineering
University of Connecticut
261 Glenbrook Road, U-37
Storrs, CT 06226-2037
wael@engr.uconn.edu

&

John N. Ivan, Associate Professor
Department of Civil and Environmental Engineering
University of Connecticut
Storrs, CT 06269-2037
E-mail: johnivan@engr.uconn.edu
Given that it is extremely unlikely that the coming years will witness major capacity-expansion projects, transportation planners will now need to view the existing infrastructure as fixed, and start thinking about how much development the current system can sustain. This line of thinking, which involves deriving land-use limits from infrastructure capacity, requires solving the inverse of the typical transportation-planning problem. While there have been some reported studies in the literature that examined how to reverse the direction of the transportation planning process, the errors of the developed models were rather significant. In the current report, we describe our efforts toward building on these earlier attempts in an effort to develop refined tools for solving the inverse of the transportation-planning problem. Specifically, our efforts have focused on developing artificial neural network (ANN) models for zonal trip ends from link volumes. The ANN models are developed and tested using real-world data from the Chittenden County region in Northwestern Vermont. Two different types of ANN models are developed, and their performance is compared. The study illustrates that ANNs are quite capable of solving the inverse of the transportation-planning problem.
DERIVING LAND-USE LIMITS AS A FUNCTION OF INFRASTRUCTURE CAPACITY

ABSTRACT

Given that it is extremely unlikely that the coming years will witness major capacity-expansion projects, transportation planners will now need to view the existing infrastructure as fixed, and start thinking about how much development the current system can sustain. This line of thinking, which involves deriving land-use limits from infrastructure capacity, requires solving the inverse of the typical transportation-planning problem. While there have been some reported studies in the literature that examined how to reverse the direction of the transportation planning process, the errors of the developed models were rather significant.

In the current report, we describe our efforts toward building on these earlier attempts in an effort to develop refined tools for solving the inverse of the transportation-planning problem. Specifically, our efforts have focused on developing artificial neural network (ANN) models for zonal trip ends from link volumes. The ANN models are developed and tested using real-world data from the Chittenden County region in Northwestern Vermont. Two different types of ANN models are developed, and their performance is compared. The study illustrates that ANNs are quite capable of solving the inverse of the transportation-planning problem.
TABLE OF CONTENTS

Technical REPORT DOCUMENTATION PAGE................................................................. ii
ABSTRACT......................................................................................................................... iii
TABLE OF CONTENTS ................................................................................................. iv
LIST OF FIGURES .......................................................................................................... v
LIST OF TABLES ............................................................................................................ v
1 INTRODUCTION AND PROJECT OBJECTIVES .................................................. 1
2 LITERATURE REVIEW ............................................................................................... 2
3 METHODOLOGY ......................................................................................................... 5
   3.1 Select Study Area ............................................................................................... 5
   3.2 Test the Transferability of the Models............................................................. 6
   3.3 Examine More Refined Modeling Techniques .............................................. 7
4 ANN RESULTS AND DISCUSSION ....................................................................... 12
   4.1 Performance of the Individual Neural Networks (Model 1)........................ 12
   4.2 Performance of the Single ANN (Model 2).................................................... 17
5 INVERSE TRANSPORTATION PLANNING: MATHEMATICAL FORMULATION .... 18
   5.1 Background ..................................................................................................... 18
   5.2 Mathematical Model Formulation ................................................................. 19
   5.3 Mathematical Model Summary ................................................................. 22
   5.4 Numerical Example and Results: ............................................................... 23
   5.5 Summary and Conclusions: ......................................................................... 25
6 CONCLUSIONS ....................................................................................................... 27
LIST OF FIGURES

Figure 1 Chittenden County Transportation Planning Model .......................................................... 6
Figure 2 The MLP Neural Network .................................................................................................. 9
Figure 3 The City of Winooski Network .......................................................................................... 10
Figure 4 ANN Prediction Accuracy for Zone 1 .............................................................................. 12
Figure 5 ANN Prediction Accuracy for Zone 2 .............................................................................. 12
Figure 6 ANN Prediction Accuracy for Zone 3 .............................................................................. 13
Figure 7 ANN Prediction Accuracy for Zone 4 .............................................................................. 13
Figure 8 ANN Prediction Accuracy for Zone 5 .............................................................................. 13
Figure 9 ANN Prediction Accuracy for Zone 6 .............................................................................. 14
Figure 10 ANN Prediction Accuracy for Zone 7 .......................................................................... 14
Figure 11 ANN Prediction Accuracy for Zone 8 .......................................................................... 14
Figure 12 ANN Prediction Accuracy for Zone 9 .......................................................................... 15
Figure 13 ANN Prediction Accuracy for Zone 10 ....................................................................... 15
Figure 14 ANN Prediction Accuracy for Zone 11 ....................................................................... 15
Figure 15 ANN Prediction Accuracy for Zone 12 ....................................................................... 16
Figure 16 ANN Prediction Accuracy for Zone 13 ....................................................................... 16
Figure 17 ANN Prediction Accuracy for Zone 14 ....................................................................... 16
Figure 18 Example Test Network ................................................................................................. 23
Figure 19 Model Performance: Objective Function and Decision Variables .............................. 24
Figure 20 Model Performance: Development Saturation Levels and Equity .............................. 25

LIST OF TABLES

Table 1 Average Absolute Prediction Error for the ANN Models ................................................... 17
DERIVING LAND-USE LIMITS AS A FUNCTION OF INFRASTRUCTURE CAPACITY

1 INTRODUCTION AND PROJECT OBJECTIVES

With the continued growth in traffic volumes, transportation professionals are faced with the challenge of how to best alleviate the nation’s ubiquitous congestion problems. Historically, congestion problems were addressed using both supply-side and demand-side initiatives. Supply-side initiatives typically included activities such as infrastructure expansion and, more recently, Intelligent Transportation Systems (ITS) projects. Demand-side initiatives, on the other hand, involve programs such as employer-based travel demand management and telecommuting. However, while these strategies have had some success in alleviating congestion, the continued growth is clearly pointing to the need for a more drastic solution to the problem.

Recently, there has been a renewed interest in better understanding and designing the land use-transportation system (LUTS), in an attempt to fight congestion more effectively. This renewed interest in the LUTS connection is not only motivated by the need to relieve congestion, but, more importantly, by the increased national interest in environmental protection and sustainability. Given this, it is extremely unlikely that the coming years will witness major capacity-expansion projects. It thus seems quite appropriate for transportation planners to start to view the existing infrastructure as more or less fixed, and to start thinking about how much development the existing system can sustain.

This new line of thinking, which involves attempting to derive land-use limits as a function of the existing capacity, requires one to solve the inverse of the typical transportation-planning problem. In the typical transportation-planning problem, one starts with a given land use pattern, and then predicts future traffic volumes on the transportation system. However, to derive land-use limits, one needs to reverse the direction of the transportation planning process; that is to say, one needs to start with the transportation system characteristics and link volumes, and use these to derive land use limits in terms of the total number of zonal trip ends. The availability of tools capable of solving the inverse of the transportation problem could be quite useful. For example, one could load the network to a desired level of service, and then use these tools to determine the corresponding land use limits that would ensure that the desired service level is attainable.

It should be noted, however, that solving the inverse of the transportation planning process is different from the so-called “integrated land-use transport (ILUT) models”, which were recently studied by the International Study Group on Land-Use Transport Interaction (ISGLUTI) under the auspices of the U.K. Transport and Road Research Laboratory (Mackett, 1993). ILUT models essentially involve linking transportation planning models to land-use models, such as EMPAL and DRAM, to allow for two-way linkages from the land use to the travel demand models and vice versa. They thus allow for modeling the long-range impacts of transportation on land-use. This is different from the inverse problem that will be considered in this study, which will focus on determining the land use pattern that will result in some sort of desired travel conditions.
To the authors’ knowledge, Miller and Demetsky were the first to attempt to reverse the direction of the transportation planning process (Miller and Demetsky, 1999). After several initial attempts, Miller and Demetsky opted for the use of a very simple, direct estimation procedure. This procedure is based on using regression analysis to develop a model that directly estimates zonal trip ends as a function of factors such as link volumes, roadway types, travel speeds, and the location of a zone relative to the other zones in the region. A second model (the socioeconomic model) was then developed to relate zonal trip to employment and population.

While the errors of the models developed by Miller and Demetsky are rather significant, their exploratory work clearly demonstrates the feasibility of reversing the direction of the transportation planning process, and helps pave the path for additional work in this area of quantifying land-use limits. It is the purpose of the current study to build upon previous research work in this area, in an effort to provide for a better understanding of the nature of reversing the direction of the transportation planning process. Specifically, the objectives of the current study are:

- To study the transferability of the relations proposed by Miller and Demetsky, using real-world transportation data sets from the State of Vermont;
- To explore the potential for including other independent variables in the proposed models, in an attempt to improve their accuracy; and
- To explore other more refined modeling techniques to develop the models, including the use of Artificial Neural Networks (ANN).

2 LITERATURE REVIEW

As previously mentioned, to the authors’ knowledge, Miller and Demetsky were among the first to attempt to reverse the direction of the transportation planning process (Miller and Demetsky, 1999). Their initial efforts to address the problem focused on attempting to solve the inverse of each of the steps in the famous 4-step planning process. Their goal was first to deduce the Origin-Destination (O-D) matrix from the link volumes, a problem which has received a lot of attention in the transportation literature, and then to deduce the zonal trip ends from the deduced O-D matrix. With the zonal trips determined, the socioeconomic characteristics for each zone can be easily deduced using regression. However, according to the researchers, this approach performed very poorly to be of use, particularly because of challenges in deducing the O-D matrix from traffic volumes.

After trying several approaches, Miller and Demetsky opted for the use of a very simple, direct estimation procedure. This procedure is based on using regression analysis to develop a model that directly estimates zonal trip ends as a function of factors such as link volumes, roadway types, travel speeds, and the location of a zone relative to the other zones in the region. A second model (the socioeconomic model) was then developed to relate zonal trip to employment and population.

Miller and Demetsky’s direct estimation procedure is based on the intuitive notion that a zone with higher volume links should be expected to have more trip ends than another zone with lower volume links, provided that the road characteristics (in terms of the ratio of through
versus local trips) of the two sets of highway segments are identical (Miller, 1998). The challenge of this approach, however, is how to discern the characteristics of the highway segments. For example, while interstate highway segments should be expected to carry high volumes, most of the volume would be a through volume and hence should not be assigned to the zones the segment is passing through. On the other hand, for local roads, most of the traffic volume could be safely assumed to originate and end within the zone. For arterials and collectors, the problem is even more challenging.

From the above discussion, it should be clear that in order to predict the number of zone trip ends from link traffic volumes, a model should be able to determine the ratio of local to through traffic on the links passing within the zone. To do this, Miller and Demetsky explored the inclusion of several exploratory variables within their model describing several elements of the transportation system characteristics. These variables can be divided in the following groups:

variables characterizing the position of a zone within the network;
variables describing the roadway type of the links contained within the zone;
variables capturing the speed characteristics of the links involved since speed could indicate whether trips are mostly through or local trips; and
variables describing count disparity between the intersecting links.

Miller and Demetsky then ran stepwise regression on these variables, using 1979 transportation data for the City of Charlottesville in the Commonwealth of Virginia. This resulted in the following model:

\[ TE = -827 + 0.027X_1 + 0.039X_2 - 1.807 \ln(X_3) + 685 \ln(X_4) + 96.1X_5 + 1729 \ln(X_6) \]

Equation (1)

where,

\( TE = \) Zonal trip ends
\( X_1 = \) number of links that are through routes (i.e. interstates & major arterials) multiplied by traffic counts
\( X_2 = \) sum of all link volumes completely within the zone
\( X_3 = \) Number of zones between the zone and the Central Business District (CBD)
\( X_4 = \) Shortest travel distance between the zone and the CBD
\( X_5 = \) Number of sampling locations used for the zone’s roadway segments (a measure of the significance of a link)
\( X_6 = \) Number of bordering zones

With the zonal trip ends estimated, Miller and Demetsky then developed a second, socioeconomic model, that related the trip ends to employment and population. This model was given as:

\[ TE = 1.64 \text{ (non-retail employment)} + 9.36 \text{ (retail employment)} + 1.11 \text{ (population)} \]

Equation (2)
The models’ accuracy was then checked using 1990 data for the same study area. According to the researchers, the Mean Absolute Error (MAE) for the combined model (i.e. the zonal trip ends model and the socioeconomic model) was about 38% of the average value. As previously mentioned, our goal in this study was to build on Miller and Demestsky’s work, in an attempt to develop better models for reversing the direction of the transportation planning process.
3 METHODOLOGY

The methodology we designed to accomplish the objectives of the project consisted of the following tasks:

Selecting the study’s test network selection and acquiring the required data;
Testing the transferability of Miller and Demetsky’s models; and
Examining more refined modeling techniques to develop the models, including the use of Artificial Neural Networks (ANN).

Each of these tasks is described in some detail below.

3.1 Select Study Area

In order to test the transferability of Miller and Demetsky’s model as well as to investigate the potential for developing more refined modeling techniques for solving the inverse of the transportation planning process, a study area was needed. For this purpose, we selected Chittenden County in Northwestern Vermont to serve as the test area for the project. The Chittenden County Metropolitan Planning Organization (CCMPO) has, since the late 1980’s, maintained an excellent transportation model for the region. This model, which consists of 325 internal zones, 16 external zones, 1,175 nodes and 1,593 links (Figure 1), was carefully calibrated in 1990 and again in 1998, and was shown to yield very accurate results. The CCMPO kindly agreed to provide the study team with all the necessary data files required for this study.
3.2 Test the Transferability of the Models

The focus in this task was on assessing the degree to which the models reported in the literature (namely Miller and Demetsky’s models) are “transferable” to other regions. As was previously mentioned above, the Miller and Demetsky’s approach is based on trying to predict the number of zone trip ends directly from the transportation system characteristics, and the sum of traffic volume on the links contained within a zone (Equation 1 above). The model, while statistically significant, had significant mean absolute errors, as previously mentioned.

One limitation of the Miller and Demetsky study, however, is the fact that the study dealt with a single study area, the Charlottesville, Virginia region. In addition, the model was based on just 43 data points, since the model utilized had only 43 zones. Given this, our objective in this task was to investigate the transferability of the Miller and Demetsky’s approach using the Chittenden County Transportation Planning Model.
County model, which contained almost seven times as many zones as the Charlottesville model. In order to do this, data had to be extracted first from the Chittenden County transportation-planning model. The data extraction effort involved looking at each of the 325 internal zones of the model, and recording the values for zonal variables that mirrored those variables used by Miller and Demetsky. The data reduction effort was facilitated through the use of ArcView GIS. Specifically, the values of the following variables were recorded for each of the 325 internal zones in the Chittenden County model:

- The number of bordering zones;
- The relative centrality of the zone with respect to the CBD (an index ranging from 1 – 10);
- The distance to the CBD in miles;
- The functional classification of the link (i.e. an interstate, arterial, collector, ..etc.);
- The number of links within the zone with more than 2 lanes/direction, which increases the likelihood that the link is used for through traffic;
- The number of links within the zone with speeds greater than 30 mph, which also increases the likelihood that the link is used for through traffic;
- Average speed on the links within the zone;
- Total and average distance of the links contained within the zone;
- The number of intersections within the zone;
- The total area of the zone in square miles;
- The sum of traffic volumes on the links contained within the zone; and
- The total number of trips produced and attracted at each zone (i.e. number of zonal trip ends).

Using the aforementioned variables, a model similar to that formulated by Miller and Demetsky was first tried. Specifically, to resemble the Miller and Demetsky’s model, the following variables were included: \( C_1, C_2, C_3, C_4, C_9 \) and \( C_{11} \). Linear regression was then run to relate the total number of trip ends to these variables, but the results obtained were rather discouraging. While the model was statistically significant, the coefficient of correlation \( R^2 \) obtained was extremely low (0.17). Efforts were then made to identify other variables and predictors from the list of the 11 variables mentioned above, in an attempt to improve the \( R^2 \) of the model. However, the best \( R^2 \) obtained was still extremely low (0.21), which shows that the linear regression approach proposed by Miller and Demetsky does not work well for the Chittenden County data. The transferability of the approach thus seems somewhat questionable, especially given the significant prediction errors and the low \( R^2 \) values for both the Charlottesville and the Chittenden County data.

### 3.3 Examine More Refined Modeling Techniques

Given the rather unsatisfactory results of the regression-based approach, the study proceeded to explore the feasibility of using Artificial Neural Networks (ANN) to model the inverse of the transportation-planning problem (i.e. predicting trip ends from link volumes). In this section, ANNs are briefly introduced first. This is followed by a description of how the data needed to train the ANNs were generated, and a discussion of the models developed.

#### 3.3.1 ANN – An Overview

...
ANNs are biologically-inspired systems consisting of a massively connected network of computational “neurons”, organized in layers. By adjusting the weights of the network, NNs can be “trained” to approximate virtually any nonlinear function to a required degree of accuracy. Starting from the pioneering work of Rosenbalt (1961), along with Minsky and Pappert (1969), NNs have evolved into what is today regarded as an important reservoir of learning methods and architectures (Rumelhart and McLelland, 1986). In general, there are two main types of learning algorithms that could be used to train a NN: supervised and unsupervised learning.

*Supervised* learning requires a *teacher* or *supervisor* to provide desired output or target output signals. This is typically accomplished by providing the network with a set of input and output exemplars (Pham and Liu, 1995). A learning algorithm (such as back propagation) would then be used to adjust the weights of the network so that the network would give the desired output. *Unsupervised* learning, on the other hand, does not require the desired output to be known. Instead, the NN is presented with input patterns, and the network adjusts the weights to cluster the input patterns into groups with similar features. Among the most-commonly cited advantages of ANNs are their ability to directly learn non-linear input-output mappings from training data, their ability to generalize to situations different from the collected data, and their ability to automatically adjust their connection weights to optimize their behavior.

Over the years, a large number of different ANNs types and architectures was developed, among the most important of which is the Multi-layer Perceptron (MLP) Neural Network. The MLP typically consists of three layers: the input layer, the hidden layer(s), and the output layer as shown in Figure 2. The type of connections in the MLP is of the feedforward type. That is to say, connections are allowed from one layer to the following layer; no connections are allowed from a layer to a preceding one or between the nodes belonging to the same layer. The MLP model was the one selected for modeling the *inverse* of the transportation-planning problem (i.e. the relationship between link volumes and trip ends).
3.3.2 **Training Data**

In order to facilitate and speedup the development and testing of the ANNs models, the study decided to use a subnetwork of the larger Chittenden County model, rather than using the whole County’s network. The subnetwork selected for this part of the study was that corresponding to the City of Winooski, a town with a population of 6,561. As can be seen from Figure 3, the Winooski network has a total of 14 zones, 45 nodes, and 114 links.
Figure 3 The City of Winooski Network

In order to generate the data required for training the ANNs, the first step was to generate a large number of O-D matrices by scaling the current O-D matrix for the City of Winooski’s network. A total of 1000 new O-D matrices were generated in this fashion, by randomly scaling the current O-D matrix, and adding noise to the values of its cells, in order to cover the solution space of the problem. The noise added to each cell was within 20% of the value of that cell. Each of these 1000 O-D matrices was then assigned to the network to yield the link volumes resulting from the assignment of that particular O-D matrix. This resulted in a total of 1000 input-output exemplars for the ANNs.

To model the inverse of the transportation-planning problem (i.e. the problem of predicting trip ends from link volumes), the input to the ANNs would have to represent the link volumes, and the desired output would have to correspond to the number of trip ends from each zone. To prepare the required training data, therefore, the number of trip ends (i.e. number of trips produced at and attracted to each zone) was first computed by summing the corresponding rows and columns from the O-D matrix (as previously mentioned, this data would represent the network desired output).

For the input data, the first step was to exclude those links that correspond to the zones’ connectors from the list of the link volumes that would be used as input to the ANN. Connectors are hypothetical links added to a transportation model in order to connect the zones to the network; they thus essentially represent the local street network in a given zone. Traffic volumes on the set of the connectors had to be excluded, because including them in the input data would make the problem of predicting trip ends from link volumes trivial (essentially the number of trip ends for a given zone is equal to the total volume on the connectors joining that zone to the network). Also, given the fact that the connectors are hypothetical links that do not correspond
to real links in the real-world transportation system, one should not expect traffic volumes to be available for these links.

After excluding the connectors and unimportant links from the network, we ended up with a total of 46 links. As mentioned above, traffic volumes on these 46 links would constitute the input to the ANNs models. That is to say, an input-output exemplar to the ANN would consist of a 46-element input vector (giving the traffic volume on each of the 46 links), and the corresponding 14-element output vector, giving the total trip ends for the 14 zones of the test network.

### 3.3.3 ANN Formulation

In this study, we have experimented with two slightly different ANN models. Both models belonged to the MLP family, and consisted of essentially three layers: an input layer, a single hidden layer and an output layer. However, for the first model (Model 1), a separate ANN was developed for each zone. This resulted in 14 individual networks in parallel, each with 46 nodes in the input layer (corresponding to the 46 network links), 3 nodes in the hidden layer, and a single node in the output layer giving the zonal trip ends for a particular zone. For Model 2, a single ANN model was developed with 46 input nodes, 3 nodes in the hidden layer, and 14 nodes in the output layer, corresponding to the zonal trip ends for the 14 zones in the City of Winooski.

The different ANN models were developed using NeuroSolutions software developed by NeuroDemension, Inc. (NeuroDimension, 2001). The previously generated data set consisting of the 1000 exemplars was divided into three groups: the training data set (60% or 600 exemplars), the cross-validation set (20%), and the test set (20%). The cross-validation set was used to test the ability of a network to generalize, while the network was still being trained. This helps safeguard against the possibility of the network “memorizing” training patterns (over-fitting), which could lead to deterioration in the ability of the network to generalize. The back-propagation algorithm was used to train the networks, and training was continued until there was no further improvement in the mean square error for the cross-validation data set during 100 epochs, or until the number of epochs reached 1000. A third data set (the testing data set) was then used to test the accuracy of the networks after training.
4 ANN RESULTS AND DISCUSSION

4.1 Performance of the Individual Neural Networks (Model 1)

Using the test dataset consisting of 200 points, plots were made of the actual versus predicted trip ends for the 14 zones in the City of Winooski’s network. The results are shown in Figures 4 through 17. These results are those obtained using Model 1, where an individual network was trained for each zone.

![Figure 4 ANN Prediction Accuracy for Zone 1](image1)

![Figure 5 ANN Prediction Accuracy for Zone 2](image2)
Predicted vs. Actual Trip Ends for Zone 3

Figure 6 ANN Prediction Accuracy for Zone 3

Predicted vs. Actual Trip Ends for Zone 4

Figure 7 ANN Prediction Accuracy for Zone 4

Predicted vs. Actual Trip Ends for Zone 5

Figure 8 ANN Prediction Accuracy for Zone 5
Figure 9 ANN Prediction Accuracy for Zone 6

Figure 10 ANN Prediction Accuracy for Zone 7

Figure 11 ANN Prediction Accuracy for Zone 8
Predicted vs. Actual Trip Ends for Zone 9

Figure 12 ANN Prediction Accuracy for Zone 9

Predicted vs. Actual Trip Ends for Zone 10

Figure 13 ANN Prediction Accuracy for Zone 10

Predicted vs. Actual Trip Ends for Zone 11

Figure 14 ANN Prediction Accuracy for Zone 11
Figure 15. ANN Prediction Accuracy for Zone 12

Predicted vs. Actual Trip Ends for Zone 12

Figure 15 ANN Prediction Accuracy for Zone 12

Predicted vs. Actual Trip Ends for Zone 13

Figure 16 ANN Prediction Accuracy for Zone 13

Predicted vs. Actual Trip Ends for Zone 14

Figure 17 ANN Prediction Accuracy for Zone 14
As can be seen from the above figures, the performance of the developed ANNs appear to be quite impressive, with the 200 points located very close to the 45-degree line (absolutely perfect prediction is indicated by all points lying exactly on the 45-degree line). In fact, for zone 1, the difference between the actual and predicted trip ends was less than 5% for 171 cases out of the total 200 test cases. That is to say, the ANN was able to predict the number of zonal trip ends to within 5% of the actual value about 86% of the time. For zone 1, the lowest absolute prediction error was 0.01%, the average was 2.81%, and the highest absolute prediction error was 9.97%. For planning purposes, this level of accuracy is quite impressive. Results for the other zones were quite comparable.

4.2 Performance of the Single ANN (Model 2)

In addition to developing individual ANNs for each zone, we also experimented with developing a single ANN model for the whole network (i.e. the 14 zones), which we refer to here as Model 2. Table 1 below compares the average absolute prediction error for the 14 zones obtained using Model 1 (the individual ANNs) and using Model 2 (the single ANN). As can be seen, Model 1 appears to outperform Model 2 for all 14 zones. By using individual ANNs, we are able to customize the model (i.e. the neural network weights) for each zone, which results in better performance.

<table>
<thead>
<tr>
<th>Zone</th>
<th>Average Absolute Prediction Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone 1</td>
<td>Model 1: 2.81% Model 2: 17.21%</td>
</tr>
<tr>
<td>Zone 2</td>
<td>Model 1: 1.69% Model 2: 17.26%</td>
</tr>
<tr>
<td>Zone 3</td>
<td>Model 1: 4.19% Model 2: 17.80%</td>
</tr>
<tr>
<td>Zone 4</td>
<td>Model 1: 2.78% Model 2: 18.66%</td>
</tr>
<tr>
<td>Zone 5</td>
<td>Model 1: 4.65% Model 2: 19.87%</td>
</tr>
<tr>
<td>Zone 6</td>
<td>Model 1: 2.02% Model 2: 18.63%</td>
</tr>
<tr>
<td>Zone 7</td>
<td>Model 1: 2.16% Model 2: 17.80%</td>
</tr>
<tr>
<td>Zone 8</td>
<td>Model 1: 0.47% Model 2: 19.44%</td>
</tr>
<tr>
<td>Zone 9</td>
<td>Model 1: 0.49% Model 2: 19.74%</td>
</tr>
<tr>
<td>Zone10</td>
<td>Model 1: 2.52% Model 2: 19.44%</td>
</tr>
<tr>
<td>Zone11</td>
<td>Model 1: 2.39% Model 2: 18.31%</td>
</tr>
<tr>
<td>Zone12</td>
<td>Model 1: 1.96% Model 2: 17.50%</td>
</tr>
<tr>
<td>Zone13</td>
<td>Model 1: 1.08% Model 2: 22.12%</td>
</tr>
<tr>
<td>Zone14</td>
<td>Model 1: 4.29% Model 2: 17.91%</td>
</tr>
</tbody>
</table>
5 INVERSE TRANSPORTATION PLANNING: MATHEMATICAL FORMULATION

In the previous sections we explored regression models and ANN for predicting trip ends based on link volumes. In part of the research, we focus on developing a mathematical formulation for the inverse or reversed transportation-planning problem. Modeling and solving the inverse transportation-planning problem will provide decision makers with insights and guidelines for determining land use policies and zoning issues.

5.1 Background

On a directed graph network $G(V,E)$, where $V$ are the vertices and $E$ are the edges comprising the graph, and where an Origin-Destination trip table is defined, the unconstrained transportation problem has been traditionally defined as:

\[
\begin{align*}
\text{Min} & \quad Z(T,G) \\
\text{SubjectTo:} & \\
\sum_{o} h_{o,d} = T_{o} & \quad \forall o \in \{O\} \\
\sum_{o} h_{o,d} \leq C_{d} & \quad \forall d \in \{D\}
\end{align*}
\]

Where,
- $Z(T,G)$ - Is an Objective function defined as a function of network traffic flow $(T)$ and the travel cost on the network $(G)$
- $P_{o}$ - Total traffic production at origin $(o)$
- $C_{d}$ - Total traffic attraction/capacity at destination $(d)$
- $h_{o,d}$ - Traffic flow between origin $(o)$ and destination $(d)$
- $\{O\}$- The set of origins
- $\{D\}$- The set of destinations

This version of the transportation problem assumes an unlimited capacity for the links in the network, which is not usually the case for urban networks. Hence, the constrained version of the transportation problem will be similar to (5-1) but with the addition of an additional set of link capacity constraints. The additional set of constraints will be,

\[
\sum_{o} \sum_{d} \delta_{i} * T_{o,d} \leq c_{i} & \quad \forall i \in \{E\}
\]

Where,
- $\delta_{i} = 1$ if the flow between origin $(o)$ to destination $(d)$ goes through link $(i)$, and $= 0$ otherwise
- $c_{i}$ - Capacity of link $(i)$
- $\{E\}$- The set of links in the graph

With the addition of this new set of link capacity constraints, the transportation problem defined in 5-1 becomes a system optimum traffic assignment problem. Wordrob (1956) stipulated the
system optimum and user equilibrium traffic assignment criteria, and suggested that the user equilibrium traffic assignment criteria to be relevant for urban network traffic assignment problems. Hence, the objective function in 5-1 for user equilibrium traffic assignment problem will be,

\[ Min \ Z(T, G) = \sum_{i} v_i \int_{0}^{t(v)} dv \]

Where,
- \( t(v) \) - Travel time function for link \((i)\)
- \( V_i \) - Total flow on link \((i)\)

Other researchers (Sheffi, et al, Daganzo, etc. Fisk) highlighted the stochastic nature of route selection and entropy in the system. Hence, the stochastic user equilibrium was introduced by modifying 5-3 by adding a term for the entropy in the system. In the literature there exists different forms for that term, here we will present the Fisk version of the model:

\[ Min \ Z(T, G) = \frac{1}{\theta} \sum_{o \in (O)} \sum_{d \in (D)} \sum_{r \in R_{od}} h_{od}^r \ln h_{od}^r + \sum_{i} \int_{0}^{t(v)} dv \]

Where,
- \( h_{od}^r \) - The traffic flow between origin \((o)\) and destination \((d)\) on route \((r)\)
- \( R_{od} \) - The set of routes between origin \((o)\) and destination \((d)\)
- \( \theta \) - A calibration parameter

The first term of this model represents an entropy component while the second term represents congestion effect. The calibration parameter \( \theta \) affects the divergence probability between alternatives; \( \theta \) can also be viewed as the parameter reflecting the error in perceiving actual travel times. This parameter in Fisk model is usually left to the users as an input. Both the Dial assignment method (Dial 1971) and the Wardrop deterministic user equilibrium assignment criterion are special cases of Fisk model. For \( \theta \rightarrow +\infty \), the first term diminishes and the model becomes a deterministic user equilibrium model. While for \( \theta \rightarrow +0 \), the first term will outweigh and diminish any effect for the second term over the value of the objective function. In this case, the model becomes a special formulation for a hypothetical case which assumes that drivers have absolute no knowledge about route costs and divert equally on all possible routes between origins and destinations.

5.2 Mathematical Model Formulation

The process of developing a mathematical model for the inverse traffic assignment problem will consist of three steps: 1) Decision variables, 2) Objective function, and 3) Constraints.

5.2.1 Decision Variables:

In this step we will identify the variables of the problem that decision makers will be interested to make decide upon. In the inverse traffic assignment problem, decision makers are usually interested in identifying the future land use pattern in the planning area in a way that will keep the transportation infrastructure near or below saturation. Land use patterns is used in the first
step in the 4-step transportation planning process to estimate trip productions and attractions, hence estimating the production and attraction by zone will give guideline for decision makers to set zoning policies that will yield this much of trips. Then, the decision variables for this problem will be:

and traffic Hence, in order to achieve this goal, the number of trip productions and attractions decision makers will be interested

- $P_i$ - Total trip productions in zone $(i)$
- $A_i$ - Total trip attractions in zone $(i)$

Notice that $P_i$ and $A_i$ encompass all trip purposes, in this formulation and for the sake of simplicity will be limited to only one trip purpose and a single mode of transportation.

### 5.2.2 Objective Function

The objective of decision makers in the inverse traffic assignment problem will be most likely to maximize the land use utility of their planning area. Hence, a weighted sum for the number of trips produced and attracted to the planning area will dictate the objective function in this context. Hence the objective function could be written as:

$$\text{Max } Z = \sum_{i \in I} wp_i \ast P_i + wa_i \ast A_i$$  \hspace{1cm} 5-5

Where,

- $wp_i$ - Weight for total production in zone $(i)$
- $wa_i$ - Weight for total attraction in zone $(i)$
- $I$ - The set of all planning zones in the planning area

The objective function defined in 5-5 could further refined to cover each trip purpose individually.

### 5.2.3 Constraints:

#### 5.2.3.1 Production and Attraction Equity Constraints:

Logically, there are upper limits for both trip production and attraction for a zone. Therefore, to ensure that the maximized trip productions and attractions at a zone will be realistic and below those upper limits, the following two sets of constraints are introduced:

$$P_i \leq P_i^{\text{Max}} \hspace{1cm} \forall i \in I$$  \hspace{1cm} 5-6

$$A_i \leq A_i^{\text{Max}} \hspace{1cm} \forall i \in I$$  \hspace{1cm} 5-7

In addition to the upper limits, equity between zones should also be considered to ensure that the growth among zones would take place within specific marginal differences. The additional equity constraints will be function of the upper limits:
\[
\left| \frac{P_i - P_j^\text{Max}}{P_i^\text{Max}} \right| \geq \beta \quad \forall i \in I, j \in \{I - i\}
\]

Where,
\[
\beta \text{-- Is an equity marginal ratio that will be set by decision makers and users}
\]

### 5.2.3.2 Trip Distribution:

In the traditional traffic assignment problem, an OD trip table was usually used as an input for the problem. Evans 1973 & 1976 combined the trip distribution part in a single combined trip distribution and assignment model. Here, we will use a logit model structure to distribute trip productions using travel time as an impedance factor.

\[
h_{ij} = P_i \times \frac{e^{-\alpha + \tau_0}}{\sum_{j \in D} e^{-\alpha + \tau_j}} \quad \forall i \in \{O\} \& j \in \{D\}
\]

Where,
- \(P_i\) - Total trip production at origin \((i)\)
- \(h_{ij}\) - Trips between origin \((i)\) and destination \((j)\)
- \(\bar{T}_{ij}\) - Average free flow travel time between origin \((i)\) and destination \((j)\)

### 5.2.3.3 Route Selection and System Entropy Constraints

Chen and Alfa (1991) have showed that the solution of the first term in Fisk model leads to a logit model structure. Hence, it will be logical to use a logit model for distributing traffic between od-pairs on a set on the set of routes connecting them. In this manner, we assure that the assigned traffic reflects the entropy in the system. The new set of

\[
h'_{ij} = h_{ij} \times \frac{e^{-\alpha + t_r}}{\sum_{i \in k} e^{-\alpha + t_i}}
\]

Where,
- \(t_r\) - Travel time on route \((r)\)
- \(v_a\) - Total traffic volume on link \((a)\)
- \(h'_{ij}\) - Traffic volume between origin \((i)\) and destination \((j)\) on route \((r)\)
- \(\delta'_{a,ij}\) - indicates if link \((a)\) is on route \((r)\) connecting between \((i)\) and \((j)\)

\[
\delta'_{a,ij} = \begin{cases} 
1 & \text{if link (a) is on route (r) connecting between (i) and (j)} \\
0 & \text{otherwise}
\end{cases}
\]

- \(v_a\) - Total traffic volume on link \((a)\)
5.2.3.4 Link Capacity and Level of Service Constraints
This set of constraints is a key factor from decision makers’ perspective. Most decision makers are in favor of expanding development in their areas, and that is reflected in the objective function 5-5. Yet, they need to achieve that goal “smartly” without creating congestion in their planning area. Hence, to assure that traffic flow on the network will not exceed capacity of the transportation infrastructure, we will include the following set of capacity constraints:

\[ v_a \leq \phi \cdot C_a \quad \forall a \in \{E\} \] 5-13

Where
\[ C_a \] – Is the capacity of link/facility (a)
\[ \phi \] – A fraction between 0 and 1 (e.g. if \( \phi \) was selected to be 0.90, then traffic volume is constrained to be 90% of capacity at most)

5.2.3.5 Adjustment for Exogenous Production and Attraction
In reality, not all productions and attractions are enclosed within the jurisdiction of the planning area. Usually, some trips will start from zones located in the planning area and end outside the planning area, and the same applies also for attraction. The amount of exogenous production and attraction depends on many factors; some are the network topology, connectivity, and regional land use. Due to the complexity of including those factors, we will assume in this part of the research that all productions and attractions are enclosed within the planning area boundaries. In fact this assumption will significantly simplify the mathematical model and eliminates the attraction parameters from the formulation.

5.3 Mathematical Model Summary
As we mentioned in the previous section, in this stage the model will be limited only to productions. The following is a summary for the mathematical model in scope of that assumption:

\[ \text{Max } Z = \sum_{i \in I} w p_i \cdot P_i \] 5-14

Subject to:
\[ P_i \leq P_i^{\text{Max}} \quad \forall i \in I \] 5-15

\[ \left| \frac{P_i}{P_i^{\text{Max}}} - \frac{P_j^{\text{Max}}}{P_j^{\text{Max}}} \right| \geq \beta \quad \forall i \in I, j \in \{I - i\} \] 5-16

\[ h_{ij} = P_i \cdot e^{-\alpha + T_q} \sum_{j \in B} e^{-\alpha + T_q} \] \quad \forall i \in \{O\} \& j \in \{D\} 5-17

\[ h_{ij}^* = h_{ij} \cdot e^{-\alpha + t_r} \sum_{i \in k} e^{-\alpha + t_r} \] 5-18

\[ t_r = \sum_{a \in r} s_a (v_a) \] 5-19
\[ v_a = \sum_{i \in A} \sum_{j \in B} \sum_{r \in R_{od}} \delta_{a,ij}^r \cdot h_{ij}^r \]

\[ \delta_{a,ij}^r = \begin{cases} 1 & \text{if link } (a) \text{ is on route } (r) \text{ connecting between } (i) \text{ and } (j) \\ 0 & \text{otherwise} \end{cases} \]

Where,

- \( P_i \) - Total trip productions in zone (i)
- \( \beta \) - Is an equity marginal ratio that will be set by decision makers and users
- \( w_p \) - Weight for total production in zone (i)
- \( I \) - The set of all planning zones in the planning area
- \( h_{ij} \) - Trips between origin (i) and destination (j)
- \( \overline{T}_{ij} \) - Average free flow travel time between origin (i) and destination (j)
- \( h_{ij}^r \) - Traffic volume between origin (i) and destination (j) on route (r)
- \( t_r \) - Travel time on route (r)
- \( v_a \) - Total traffic volume on link (a)

5.4 Numerical Example and Results:

5.4.1 Test Network

The model developed in this part was tested on a simple network to validate the approach and to check for unsoundness. The test network consists of four links and two origins and two destinations as shown in Figure 18. The attraction at destinations was assumed to be unlimited.

![Figure 18 Example Test Network](image)

5.4.2 Model Solution Method:

The following are the steps we followed for solving the suggested mathematical model:
**Step 0:** In this step we assumed an empty network and that link travel times are the free flow travel time $t(0)$.

**Step 1:** Using network travel time from the previous iteration, we used MS Excel Solver™ to maximize the objective function.

**Step 2:** Check for Convergence: In this step we compare the value of the objective function at the current iteration ($n$) with the value at iteration ($n-1$), the stopping criterion was:

$$\left| \frac{Z_n - Z_{n-1}}{Z_n} \right| \leq e$$

Where $(e)$ is a convergence factor selected by the user (in our case we used $e=0.001$)

**Step 3:** If the stopping criterion in Step 2 was not met, update network travel times and go to step 1, otherwise stop

### 5.4.3 Results and Discussion:

#### 5.4.3.1 Objective function and model Convergence:

Figure 19 shows the values of the objective function ($Z$) and the productions in zones 1 and 2 (P1&P2). It could be observed from Figure 2 that the model was conversing fast, that was expected especially with such a very simple network. Despite the smallness of the network in this example, we confidently can say that the model is bounded and has an optimum solution.

![Figure 19 Model Performance: Objective Function and Decision Variables](image)
5.4.3.2 Equity and Development Saturation:

Figure 20 illustrates the ratios between productions and production upper limits for both origin 1 and 2. These ratios are Sat1% and Sat2% reflecting the developmental saturation levels in each zone. In addition the marginal difference between the developmental saturation levels for those zones are represented by the equity ratio. It is clear that in this example, the network topology and developmental capacity for zone 1 favored it over the other zone and granted it more development opportunity than zone 2.

![Figure 20 Model Performance: Development Saturation Levels and Equity](image)

5.5 Summary and Conclusions:

In this part of the research we presented a mathematical model for the Inverse Transportation Planning Problem. The developed model reverse the 4-step planning process and make land use as the decision variables in the problem. The current model is subject to the assumption that all productions and attractions are contained within the planning area and ignores traffic passing through the planning area. These assumptions will be subject to future modifications of the model. In addition a efficient solution method for the model needs to be developed. However, the developed model showed to be bounded and conversed on a small size network.
In conclusion, the developed model has a strong potential for rethinking the transportation planning and zoning process. In addition, it raises other considerations such as equity between zones in the planning area.
6 CONCLUSIONS

Based on results from this study, the following conclusions can be made:

(1) The transferability of the regression-based approach developed by Miller and Demetsky for predicting land use limits from infrastructure capacity appears to be questionable.

(2) ANNs are quite capable of mapping the relationship between network link volumes and zonal trip ends, and hence can be used to predict land-use limits from infrastructure capacity. The training data can be easily generated from an existing transportation planning model for the region.

(3) The ANN model could be regarded as “transferable” from one region to another (in the sense that the ANN architecture need not be changed). However, the network will have to be retrained using data generated from the new region’s specific transportation-planning model.

(4) The ANN approach can be used to solve the inverse of the transportation-planning problem for networks larger than the one considered in this study, provided that enough training data is generated.

(5) For solving the inverse of the transportation-planning problem, the training of individual ANNs for each zone is preferable over training a single ANN for the whole network.
REFERENCES


